Identifying Regional Industry Clusters in California: An Initiative of EDD's Labor Market Information Division (LMID)

A Summary

April 2001

Employment Development Department
State of California

LMID
Labor Market Information Division
"Your Information Source"

Identifying Regional Economic Clusters in California

What Is an Industry Cluster?

An industry cluster is a subset of industries in the regional economy connected by flows of goods and services stronger than those linking them to the rest of the economy. Individual firms in a cluster benefit from certain comparative advantages associated with geographical concentration such as access to a common pool of specialized labor, infrastructure, intellectual property, and lower transaction costs between firms.

Labor Market Information Division's (LMID's) Industry Cluster Initiative.

♦ Background.

Industry cluster analysis has generated considerable interest in the development community as a tool for designing effective regional economic development strategies. LMID first became involved in industry cluster analysis in its support role to the California Economic Strategy Panel in 1994. The Panel conducted an analysis of the industrial clusters in ten California regions based on information provided by local economic development representatives in each region.

A wide variety of quantitative industry cluster definitions and identification models have been proposed in the regional economic growth literature, but none of them have stood out or received widespread acceptance. In late 1998, in response to a request for a generally accepted, data-based method to classify industry clusters, LMID issued a Request For Proposals (RFP) for the "development or identification of a practicable model(s) and methodology for defining and studying industry clusters in any area or region of the state. The model should be standardized to the degree that economic development specialists and planners throughout the state use a common language and methodology; yet flexible enough to endure the dynamics of the California economy."

Forty-eight potential bidders, representing 42 different organizations, expressed interest in the RFP. Two proposals were submitted. A panel comprised of LMID staff and outside experts evaluated the relative merits of each proposal.

The winning proposal was submitted by San Diego State University. Subsequently, a contract between EDD and the San Diego State University Foundation was signed in 1999, and the project was completed in July 2000. Dr. Sergio Rey, Associate Professor, Department of Geography, San Diego State University (SDSU), an expert in the fields of integrated multi-regional modeling and forecasting, methods of regional economic analysis, economic geography and spatial econometrics, was the Principal Investigator on this project. He was assisted by Mr. Daniel Mattheis, a graduate research assistant in the SDSU

Department of Geography. An evaluation team comprised of representatives from two local agencies and a private economic developer helped guide the research.

The research resulted in four report volumes: 1) Conceptual Design, 2) Methods Handbook, 3) Technical Documentation of Statewide Industry Clusters, and 4) The Role of Industry Clusters in California's Recent Economic Expansion.

- **What Was Done.**

The study included a thorough evaluation of the quantitative cluster identification models in the literature. These were classed into two general analytical approaches, with each having multiple specifications (or variations).[1] Over 170 variations were tested using an input-output model (IMPLAN 1997) for the State of California. The findings confirmed that no model performed better than others in identifying industry clusters. Each analytical approach tended to stress certain inter-industry relationships (or links), but ignored others. None of the analytical approaches or variations could capture the complexity of economic relationships between industries in a cluster, including both internal relationships (forward, backward, and complementary links) and external links to the larger regional economy. Also, the industries included in any given cluster tended to be sensitive to choice of technique and measures of relationship strength between industries.

A New Cluster Identification Method: The Consensus Clustering Approach

- **The Glass Is "Half Full".**

The consensus clustering approach proposed by Rey and Mattheis integrates the various cluster identification models into a single approach. Its premise is that while existing cluster identification methods may have limitations that compromise their usefulness, they nevertheless produce important information about the economic relationships between industries in a cluster. The consensus clustering approach identifies pairs of industries that join consistently together across the various methods. By doing this, the method considers multiple types of economic relationships (forward, backward, external, and complementary) between firms in defining clusters.

The consensus clustering approach consists of two main steps. First, programs containing the 81 most informative of the 170-plus variations of cluster identification methods are run to generate individual cluster profiles, or lists of industries that are grouped together to form clusters. Second, each of the individual cluster profiles obtained in Step 1 are used as inputs to generate a single "consensus cluster" profile. This process entails constructing a

[1] More detailed descriptions of these approaches are found in the Appendix.

dissimilarity matrix that measures the extent to which any given pair of industries were not placed in the same cluster by the 81 different methods. The greater the linkages between industry pairs across methods, the lower is the dissimilarity value.

The results may be presented in a dendrogram, which is a graphical depiction of cluster/industry relationships in the form of a tree diagram. Trimming the diagram determines which branches of industries remain together as a cluster. This is done by drawing a horizontal line at different "dissimilarity values" which measure the strength of relationships between industry pairs. If the line cuts a branch, these industries are included in a cluster. If the dendrogram is trimmed at a low enough value, no industries will form a cluster. If trimmed at a high enough value, all industry pairs join to form a single cluster. Trimming the dendrogram at different intermediate values generates alternative cluster profiles that can be examined for content and consistency (i.e., Are there too many or too few clusters? Is there a single large cluster and several small ones? Do the industries within the cluster make sense economically?) Because local knowledge and expertise play a key role in guiding the ultimate definition of final industry clusters, the consensus clustering combines qualitative and quantitative criteria in defining industry clusters.

The consensus clustering method uses two open source software packages--Octave and R. A methods handbook has been developed that provides explicit, step-by-step data and diagnostic programs to guide economic development practitioners with differing levels of technical expertise in applying the method.

Statewide Findings.

The consensus clustering method was used to identify industry clusters using industry-level data from the IMPLAN input-output model for California. Of the 515 industries in the model, 149 were identified as potential cluster members. These were industries with a strong export orientation as indicated by their having location quotients exceeding one.[2]

- The method identified 35 multi-industry clusters in California that included a total of 114 industries (See Table 1 and 2, following pages). Thirty-five industries did not cluster and were assigned to the non-cluster category.

- The largest of the consensus clusters, General Manufacturing, included 21 inter-related industries. Twenty-five clusters consisted of industry pairs only.

[2] Location quotients are a generally accepted measure of an industry's export orientation. They are derived by dividing the share of an industry's employment in total regional employment by that same industry's share of employment in total national employment.

- The 35 consensus clusters span most economic sectors and production chains in the California economy.

Consensus Cluster Characteristics.

- The 35 consensus clusters accounted for 45% of California's total economic production in 1997. The largest clusters were Real Estate ($129 billion in output), Wholesale Trade and Miscellaneous Business Services ($118 billion), and Computer and Electronics Manufacturing ($92 billion).

- Consensus clusters accounted for 39% of the California's total employment in 1997. The largest clusters were Retail Stores (over 1.3 million jobs) and Wholesale Trade and Miscellaneous Business Services (over 1 million jobs). Services and trade clusters tended to dominate by employment size.

- Cluster industries tended to have higher than average labor productivity, as indicated by their share of total output being greater than their share of total employment.

- Wage levels in the clusters varied substantially. Average wages in the top four clusters were over twice the statewide average, but only 60% of the State average in the bottom four clusters.

- Only the Computer and Electronics Manufacturing cluster ranked in the top-ten among clusters with respect to employment and wages.

- Consensus clusters displayed very strong external links with the rest of the California economy. Thirty-one of the 35 clusters displayed above average supply (forward) or purchase (backward) linkages. Clusters with strong purchase linkages can be considered to be drivers of the state economy, while clusters with strong supply linkages can be considered to be enablers of expansion.

Table 1
Multi-Industry Consensus Clusters in California

Food Sector

- Dairy
- Wine
- Vegetables, Cotton, and Other Crops
- Fruits
- Nuts
- Food Canning
- Miscellaneous Food Manufacturing

Mining Sector

- Sand and Gravel
- Petroleum Refining and Gas Production

Manufacturing Sector

- General Manufacturing
- Office Equipment Manufacturing
- Medical Equipment and Supplies
- Household Entertainment Appliances
- Computer and Electronics Manufacturing
- Analytical Instruments
- Missile, Aircraft, Navigation Equipment
- Automotive Trimmings

Trade Sector

- Retail Stores
- Wholesale and Miscellaneous Business

Services Sector

- Miscellaneous Personal Services
- Repair Services
- Office Machine Services
- Advertising
- Religious Organizations
- Miscellaneous Services
- Personnel Services
- Non-Profit and Educational Services
- Business and Rental Services
- Motion Pictures and Broadcasting
- Race Tracks
- Entertainment

Finance, Insurance and Real Estate Sector

- Credit Agencies
- Real Estate
- Insurance and Management Services

Government Sector

- Local Utilities and Services

Table 2
Industry Composition of Selected Consensus Clusters

Computer and Electronics Manufacturing Cluster

- Electronic Computers
- Computer Storage Devices
- Computer Peripheral Equipment
- Telephone and Telegraph Apparatus
- Radio and TV Communication Equipment
- Communications Equipment N.E.C.
- Printed Circuit Boards
- Semiconductors and Related Devices
- Electronics Components N.E.C.
- Electrical Equipment N.E.C.
- Mechanical Measuring Devices
- Instruments to Measure Electricity
- Electromedical Apparatus

Retail Stores

- Food Stores
- Apparel & Accessory Stores
- Furniture & Home Furnishings Stores
- Miscellaneous Retail

Local Utilities and Services

- Water Supply and Sewerage Systems
- State and Local Electric Utilities
- Other State and Local Government Enterprises

Motion Pictures and Broadcasting

- Radio and TV Broadcasting
- Motion Pictures

Role of Clusters in California's Economic Recovery 1992-1997.

Consensus clusters played a pivotal role in the California economy's post recession expansion between 1992 and 1997.

- Consensus clusters generated 53% of the 1.8 million net jobs created in California between 1992 and 1997, despite their having accounted for only 37% of total employment in 1992. In contrast, the larger non-cluster segment of the economy accounted for only 47% of net jobs created.

- The average annual rate of employment growth in clusters was 3%, compared to only growth of only 1.7% in the non-cluster segment of the economy and 2.2% in the State overall.

- Five industry clusters each added over 100,000 new jobs between 1992 and 1997. These were: Retail Stores, Personnel Services, Business and Rental Services, Wholesale and Miscellaneous Business Services, and Credit Agencies.

- Twenty of the 35 consensus clusters had annual employment growth rates above the State average of 2.2%.

- Eleven clusters expanded at the national level and grew faster in California than elsewhere in the country. Seven of these also ranked in the top-ten among consensus clusters in net job creation. They were: Retail Stores, Business and Rental Services, Insurance and Management Services, Motion Pictures and Broadcasting, Non-Profit and Educational Services, Computer and Electronics Manufacturing, and General Manufacturing. These clusters were key forces in driving California's post-recession economic recovery.

Current Research.

The research study results demonstrated that the consensus clustering method holds considerable promise as a useful tool for regional economic development programs. But further investigation is necessary before this potential becomes fully realized. A follow up research agreement between EDD and San Diego State University was drawn and signed in early March 2001. The research, now underway, will:

- Evaluate how the consensus clustering method performs at the regional, or sub-state, level. Four major test regions within the State will be used to implement and evaluate how the model identifies regional specialization.

  1) San Diego: San Diego County

2) Southern California: Imperial, Los Angeles, Orange, Riverside, San Luis Obispo, Santa Barbara, San Bernardino, and Ventura Counties.

3) Bay Area: Alameda, Contra Costa, Marin, Monterey, Napa, San Benito, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano, and Sonoma Counties.

4) Central Valley: Fresno, Kern, Kings, Madera, Merced, San Joaquin, Stanislaus, and Tulare Counties.

- Identify the key decisions that analysts must make when implementing the method and examine how these decisions affect its results--especially at the regional and local level. Specifically, the study will focus on how to determine export base thresholds, dendrogram trimming, assigning secondary industries to clusters , and how to incorporate qualitatively produced cluster definitions into the analysis

- Further refine the consensus clustering method to enhance its use by economic development practitioners.

Appendix

Why Are There So Many Cluster Identification Models?

The quantitative cluster identification methods proposed in the literature seek to assign individual industries to an industry cluster based on the nature and strength of their interrelationships. Input-output models, which provide a detailed accounting of purchases and sales between industries (firms), form the core of these methods. Rey and Mattheis evaluated 171 cluster identification implementations, which they grouped into two general classes of approaches

The Principal Components Factor Analysis (PCFA) approach identifies elements of a cluster by analyzing the principal components of inter-industry flows. PCFA implementations vary with respect to how the input matrix used to isolate these flows is manipulated and standardized, and the "factor loading thresholds" used to measure the strength of inter-industry relationships within the cluster. Rey and Mattheis evaluated 36 PCFA models (six alternative specifications of input matrices using six different factor loading thresholds).

Strengths of PCFA Methods:

- Highlight external and complementary linkages between industries

Weaknesses:

- Poor in identifying vertical linkages and commonalities between firms.

The Multivariate Clustering (MVC) approach uses algorithms to group industries together such that members of one cluster are similar to each other yet distinct from industries outside the cluster. This entails specifying an objective function which establishes the criteria used to group industries together; developing a measure of the similarity (or dissimilarity) between industries in the group, and sequentially clustering industries based on the strength of their relationships. Rey and Mattheis evaluated 135 variations of MVC models (three "hierarchical" algorithms times five alternative input matrix specifications times 8 strength of relation thresholds plus 15 iterations of a "partitioning" algorithm).

Strengths of MVC Methods:

- Highlight internal (vertical) linkages between firms.
- Certain algorithms produce a dendrogram, which provides a visual picture of the links between firms.

Weaknesses:

- Poor in identifying complementary linkages and chaining effects.